DOCUMENT RESUME

ED 165 634 HE 010 898

AUTHOR North, Joan; Scholl, Stephen
TITLE Revising a Faculty Evaluation System. A Workbook for Decision-Makers.
PUB DATE 78
NOTE 61p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Administrator Education; Administrator Guides; College Administration; *College Faculty; Contracts; Data Collection; Decision Making; *Evaluation Methods; *Faculty Evaluation; Faculty Workload; Higher Education; Instructional Materials; Measurement Instruments; Measurement Techniques; Program Development; *Program Evaluation; Student Evaluation of Teacher Performance; *Teacher Evaluation; University Administration; Workbooks; Workshops
IDENTIFIERS Growth Contracts

ABSTRACT

This workbook, designed for use in a workshop or independently by a college or university group, is intended for use by faculty and administrators reviewing an existing faculty evaluation system. The tool is organized in chapters outlining sequential steps in decision-making about an existing evaluation system. Following each chapter is a series of group discussion questions. The chapters address these issues: reviewing an evaluation system; evaluation philosophy and the nature of the college; purposes of faculty evaluation; identifying and weighting faculty activities; using data to support evaluation; schedule and responsibility for collecting evaluation data; and instruments and procedures, adoption, implementation, evaluation, and further review. Appendices include Using Student Appraisal Instruments; A Case for Differential Evaluation; and Growth Contracts. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED165634

REVISING A FACULTY EVALUATION SYSTEM

A Workbook for Decision-Makers

Joan North
Stephen Scholl

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Joan North

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HE 010 898

REVISING A

FACULTY EVALUATION SYSTEM

A
Workbook
for
Decision-Makers

Joan North
Stephen Scholl

© Copyright 1978 Joan North and
Stephen Scholl

CONTENTS

I. Introduction 2

II. Reviewing an Evaluation System 5

III. Evaluation Philosophy and the Nature of the College. 10

IV. Purposes of Faculty Evaluation 15

V. Identifying and Weighting Faculty Activities 18

VI. Using Data to Support Evaluation 19

VII. Schedule and Responsibility for Collecting Evaluation Data 22

VIII. Instruments and Procedures, Adoption, Implementation, Evaluation and Further Review 23

General Bibliography. 25

Appendices

1. Using Student Appraisal Instruments 26

2. A Case for Differential Evaluation. 36

3. Growth Contracts. 39

I.

INTRODUCTION

This workbook is designed for use in a workshop or independently by a college or university group organized to revise a faculty evaluation system. We presume that, unless the particular institution was organized yesterday, some form of evaluation already exists. It may not be satisfactory or it may be quite good, needing only slight revision. You may find the individual chapters, forms and appendices that follow more or less helpful. We trust that you will make judgments about how much time you need to spend on each section and what issues are most critical for your setting.

We have written this workbook primarily for faculty and administrators who will be making decisions about the design and implementation of an evaluation system. Many colleges and universities are discovering a need for more formal evaluation, greater emphasis on institutional rather than individual professional needs and public accountability of those who make personnel decisions. Some of the factors that demand better evaluation systems are unfortunate: lessened faculty mobility, enrollment decreases and financial stringency. Other signs are more encouraging: an interest in more humane and fair reward systems and the increasing popularity of practical and effective faculty development programs. Hopefully, your system will be revised both because it needs to be and because you believe the institution will be a happier and more productive place to work with a better evaluation program.

The workbook is divided into chapters which outline sequential steps to making decisions about a system for faculty evaluation. Following

each chapter is a series of questions to be answered by a group within a college or university before proceeding to the next stage. These questions may serve as the backbone of a weekend workshop on faculty evaluation or may be used throughout the process of revising a faculty evaluation system.

The authors of this workbook make some basic assumptions about evaluation. We believe that the evaluation system must be compatible with the formally stated college mission and the operational style of the institution. We assert that there should be public discussion of the philosophy and purposes of evaluation. We believe that evaluation systems need to be both thorough and sensitive, but also easy to maintain. We believe that those who are to be evaluated must be involved in the design, implementation and review of the process. We think that everyone who is served by faculty should have a hand in evaluation -- students, of course, but also colleagues, the administration and significant others, perhaps including Trustees and some people who are external to the institution. On the other hand, we think there is no need and little justification for proceeding as though no one else had ever addressed the evaluation problem before. Any system must fit your peculiar situation; it should also be based on the best research and development that has gone on elsewhere. Finally, because evaluation is such a complex process, we think that no one approach or method is perfect or adequate. Multiple sources of data and a variety of approaches to making evaluation decisions are the greatest safeguard in designing an effective system.

Based on the literature and our own experiences, we have tried to suggest an orderly process for reviewing and revising a faculty evaluation system. We have consciously skipped over issues that do not seem to us

to be at the center of the review process. However, we recognize that some of these popular topics will need to be discussed in any workshop or by any group which approaches evaluation redesign. Therefore, in the appendices we have included sections on student rating instruments, differential weighting of faculty activities and growth contracts. As these issues arise, the appendices may prove to be helpful guides for discussion. However, these topics themselves are not as critical as is an orderly review of the process of revising an evaluation system. They are merely elaborations on techniques that may need to be considered.

We do not presume that any group proceeding through this workbook will end up with an evaluation system proposal identical to that produced by any other group. But the differences should reflect the differences between institutions, not a lack of consideration of generic issues. Completion of this workbook and the redesign of a system should be seen as one stage in the development of an evaluation process. Continual review and revision is what will keep the evaluation system alive and well.

Joan North
Director of Faculty and Staff Development
Small College Consortium
Washington, D.C.

Stephen Scholl
Dean of Educational Services
Ohio Wesleyan University

II.

REVIEWING AN EVALUATION SYSTEM

A helpful approach to revising an existing faculty evaluation process is to assume and act as if you are beginning from scratch. Ask basic questions, such as: Who is asking for faculty evaluation? Who will want to see the results of evaluation? Who is interested in the process? The initiatives to review a program (or begin one in the case where none exists) may emanate from existing representative groups, such as faculty, students, the administration, even parents, trustees or legislators. Who is interested and who they represent are important issues. If there is not some agreement among a significant set of groups that something needs to be done or changed, there is little point in proceeding any further.

Once the decision to review has been made, the process of revising an evaluation program can be divided into phases, focusing in turn on study, adoption, implementation, evaluation and review once again. In this chapter we will focus most of our attention on the study phase.

Study Phase

The formation of a well-structured study group is critical. It should be neither too small nor too large; six-to-nine participants is workable. Although those to be evaluated must be the most significant participants, at least three constituencies are usually needed in a typical college or university -- faculty, administration and students. Sometimes members of an institution's governing board may also be interested in working with this group.

The study group's objectives should include high visibility, open and broad communication, credibility (who is on the committee and how they are selected are the most important initial factors in establishing credibility), the gathering of extensive data and thorough understanding (reading, workshops, etc.) of important evaluation issues.

Most study groups work more effectively if they use a standard group problem-solving process.[1] The following is a typical process:

1. Define and get agreement within the group on the problem (spend sufficient time at this stage or all that follows will be inefficient).
2. Set goals (statements that are the reverse of the identified problem) -- what should the situation be when the process is completed?
3. Develop and consider several alternative solutions, including, perhaps, maintenance of the status quo.
4. Analyze the forces for and against change (technically, "force-field analysis") on paper and concentrate attention on forces that resist change rather than those which promote change (the latter are already on your side).

---

[1] For more detailed guidance on group problem solving see William Morris and Marshall Sashkin, "Phases of Integrated Problem Solving," in Pfeifer and Jones, eds., The 1978 Annual Handbook for Group Facilitators (La Jolla, Cal.: University Associates, Inc., 1978), pp. 105-116. For alternate techniques, see Arthur Chickering, et al, Developing the College Curriculum, Appendix C ("Planning Tools"), (Washington, D.C.: Council for the Advancement of Small Colleges, 1977), 279-290.

5. Plan a strategy, including suggestions regarding what to do, by whom and when. Put the strategy in draft form and distribute it among the major constituencies of the institution, perhaps separating substantive issues from concerns with the evaluation process.
6. Provide for at least one public campus discussion of revision proposals at this "draft" stage. At such a meeting strategy itself is important -- for instance, a good visual model, such as an overhead projection or large chart focuses the attention of the group up front, together, and indicates a common search for solutions rather than tangential critiques of minor aspects of the process.
7. Make recommendations, incorporating suggestions from the campus meeting and including means for periodic revision of the system.

Any small group that is charged with wrestling through the problems of evaluation is likely to face some tensions in its own functioning. The group is in the spotlight; it must produce recommendations that directly impact on the livelihood and well being of many people. Further, academic committees are not famous for efficiency, teamwork or decisive action. So the study group ought to attend to its internal process of decision-making and taking action. Technical advice abounds on small-group effectiveness.[2] At minimum, the group should begin by deciding how

---

[2] Some examples of straightforward primers on effective small group work are: "Decision-Making," in William Bergquist and Steven Phillips, A Handbook for Faculty Development, Vol. I (Washington, D.C.: Council for the Advancement of Small Colleges, 1975), pp. 157-164; and Andre Delbecq, Andrew H. Van de Ven and David H. Gustafson, Group Techniques for Program Planning (Glenview, Ill.: Scott, Foresman and Co., 1975).

to decide (voting, consensus, or some interactive process) and how to act (by delegation through a staff, by sharing assignments, through a chairperson, etc.).

The entire study and recommendation process will take about five-to-nine months under normal circumstances, with implementation usually coming in the following academic year. If the study process itself takes more than a single academic year, energy for the issue tends to wane and people begin to believe that nothing will happen. A fairly tight time schedule, with special meetings for discussion of the revision of the evaluation program, is much to be desired.

There are several common sense details to be observed during the study process. For instance, it is wise for the study group to share its thoughts with smaller groups of faculty informally prior to circulating a general draft proposal and having a large campus meeting. It is usually advisable not to have the general campus meeting at a regular faculty meeting, to avoid the notion that an early decision is going to be forced and in order to have it truly perceived as a time for study and suggestions rather than legislation. It is important that the full study group participate in any campus meetings to show general support for the proposals and to be available to answer any specific suggestions that may represent differing perspectives. Such participation continues to enhance credibility for the study process.

Other Phases

During the adoption phase those to be evaluated, the faculty, should be able to thoroughly analyze recommendations and formally accept ownership of the evaluation system. If they cannot be convinced that

this is the best way to go for them, it is better to revise recommendations than to force an unacceptable program. During the implementation phase, the critical element is a smoothly functioning process that is perceived as thorough and yet does not take so much time that it becomes burdensome. Again, careful planning and execution are important. The evaluation phase is a continuous tribute to a belief in evaluation of all important enterprises in the institution. There is always something that can be improved in any process. Finally, it simply must be assumed that any new program may in turn be revised again, so building in a time for formal review is only good planning.

In Form 1, which follows, are a series of questions pertaining to the creation of the study process. The answers to these questions should provide you with the basic action planning design for reviewing faculty evaluations.

FORM 1

Determining the Study Process

1. Who initiated this review?

2. Who will the study group be? How is it to be chosen? What groups and how many from each group are to be representative?

3. To whom is the study group responsible?

4. How will the study group keep the faculty informed?

5. How will the study group operate in making decisions and taking action?

6. How will a final decision to adopt and implement a revised evaluation process be made?

7. What is the general list of activities for the study and recommendation process?

8. Chart the important steps to be used in the study process; give completion deadlines.

9. Consider using this completed work sheet for a check on the process midway through the study (especially just prior to any general campus meeting called to review draft recommendations). The study group should ask itself: Has anything changed since we began? Are we on schedule? If there are revisions or problems in the process, are they being appropriately dealt with? Are we functioning effectively as a group? Finally, review this sheet just prior to proposing a revised evaluation process for adoption. Ask the same questions that were asked at the midpoint. Checking the process in this way can prevent unfortunate surprises at critical stages.

III.

EVALUATION PHILOSOPHY AND THE NATURE OF THE COLLEGE

Although few people pause to examine an evaluation philosophy, the consideration is helpful for ensuring congruence between evaluation practices and the nature of a particular institution. If the system of evaluation is incompatible with the college mission and style, it will not be operational. Imagine, for instance, the difficulties for a small college which emphasizes community and individual growth for its faculty and students while it employs a competitive, comparative evaluation system.

This chapter outlines four generalized evaluation philosophies. The approaches overlap and they are admittedly ideal types. Discussion of these types may help to clarify a study group's philosophical approach to the problem. The next chapter will focus on more specific purposes or functions of evaluation.

1. Summative comparison with other faculty. Comparisons can be made within the college, within a department, with faculty nationally, within a discipline nationally, with faculty of the same rank, or with faculty having similar classes in the school or nationally. Most faculty evaluation systems have traditionally relied on such comparisons and resembled periodic "exams", with the tenure decision perceived as the "final". It is understandable that the faculty member should ask, "With whom am I being compared, and are there any differences in our situations which may work against me?" Summative comparisons often employ quantification, but this approach does not require it.

Advantages: It is sometimes easier to judge people in comparison to other people than to measure them against an ideal too often not clearly defined, such as "effective teaching." So, this approach can provide relatively objective comparisons and can help to make difficult administrative decisions by relying on differences between performance and production of various individuals. When resources are scarce, there can be a ranked list of people to be rewarded, promoted or dismissed. One ends up with the cream of the crop.

Disadvantages: The discrimination between individuals may appear to be so slight that decisions are not defensible. There is certainty only about who are the best and worst faculty. This type of summative evaluation generally comes at the end of a semester or year, and the burden of proof is on the instructor. Individuals may be compared to inappropriate groups, for example, to someone in an institution with more support for research or to someone with smaller classes in a more popular subject. The approach forces a ranking, and implies that some faculty must be performing poorly even if the differences are slight. The bell curve is not always appropriate.

2. Criterion-based evaluation. This philosophy assumes that there is an ideal standard or a set of basic criteria against which faculty are judged. When the criteria are explicit, they may include a fixed number of publications or a certain score on student evaluations or other performance measures. Often, however, the ideal is only implicit in personnel decisions and not explicitly publicized.

The "ideal standard" may be a "Mr. Chips" type, the internationally renowned scholar, or the unofficial composite of favored tenured professors.

Advantages: If used explicitly, this philosophy promotes discussion about ideals or criteria which may result in useful institutional standards and can, therefore, provide for quality control. Explicit criteria also provide clear guidelines for faculty who are able and willing to mold their behavior to fit the standard.

Disadvantages: This philosophy may limit tolerance for individual differences. The philosophy may lead to a trait approach to evaluation, rather than looking at behavior or student learning outcomes. Thus it may have a tendency to promote "looking right" and effusive game-playing. If, on the other hand, the criteria are not explicit, faculty may be evaluated on evidence about which they are ignorant.

3. Developmental or formative evaluation for growth. This approach emphasizes the individual and personal growth, rather than a relative ranking compared to others or to an ideal standard. A "growth plan" philosophy implies that everyone can and should improve in performance and the change, rather than the relative strength, is what is rewarded. Colleges recruit good faculty who show promise of continued growth. Faculty should understand clearly and accept what needs to be done during the evaluation period.

Advantages: Comparison is with oneself and is related directly to one's professional growth. Evaluation frequently begins at the start of a semester, year or several-year period

allowing time to change and improve. This approach may stimulate more productive effort than the summative comparative approach and may produce more behavior change. Development and evaluation may be combined in this approach. Multiple sources of data are clearly required. The process is highly individualized.

Disadvantages: Developmental evaluation takes more time because each person must be dealt with at length, although not necessarily every term or year. If an individual's personal growth plan is to benefit the department or college, considerable customizing and planning are needed early in the process. It may be difficult to establish and maintain equitable individual expectations across an entire institution. Even with improvement, a person may still be performing so poorly that the college wishes to terminate him, but the developmental approach has not been used often to support dismissal. There can be an implied "contractual" obligation to retain everyone, though institutional needs may change during the period of evaluation.

4. Objectives-based evaluation. This approach is borrowed from the military and business world, where it is usually labeled "Management by Objectives". This philosophy holds that the creation of annual individual measurable objectives will lead to explicit results. To be effective, the entire institution must be involved, from top to bottom. This philosophy essentially suggests a contract approach that "if fulfilled, results in some individual reward." The usual difference between this philosophy and the developmental approach is that the objectives established for individuals are

based primarily on the institution's needs, not the individual's. Results, rather than process, are most valued.

Advantages: This philosophy clearly establishes the primacy of the institution over individuals within the college. It requires that individuals agree on institutional objectives and their own responsibilities for helping to achieve them. It forces the college to be clear about its objectives. It is systematic, reasonably predictable and may provide a unifying force within the institution.

Disadvantages: This philosophy involves the creation of periodic goal-setting and review sessions with all college personnel. Thus, it is time-consuming. Because this approach is so much tied to institutional objectives, it may ignore individual needs or stifle creativity. Also, it may not be possible to objectify the most valued goals of institutional or individual life. Not every faculty activity can be derived from the college mission statement.

The form following this chapter asks you to choose which evaluation philosophy most closely matches your institution.

FORM 2

Evaluation Philosophy and the College Mission

Consider your college's mission (formally stated), general operating style and attitude toward people. Based on what you perceive, select the evaluation philosophy which you believe should be most appropriate for your college, and explain why you believe it fits. Suggest any important modifications in the approach which might be necessary.

1. Effective faculty evaluation compares a person's performance with others' performance to determine who is performing better at critical junctures (summative comparisons).

2. Effective evaluation measures a person's performance against specific criteria or an ideal standard (criterion-based evaluation).

3. Effective faculty evaluation measures a person's improvement in performance periodically (developmental/formative evaluation).

4. Effective evaluation measures regularly the results a person achieves against predetermined objectives (objectives-based evaluation).

5. Other.

IV.

PURPOSES OF FACULTY EVALUATION

No matter what the institutional philosophy, most evaluations have at least two implicit purposes -- to make judgments about an activity so it may be appropriately valued and rewards may be distributed, and to improve the activity itself. One of the basic problems about faculty evaluation systems is confusion about using evaluation for personnel decisions and/or for faculty development. While evaluation data is almost always used for personnel decisions, the types of information gathered may be more appropriate for developmental purposes. For instance, many student evaluations of instruction contain detailed questions about specific teacher behaviors as well as summary statements about the overall effectiveness of a faculty member and a course. Probably only the summary statements about the teacher are useful in making key personnel decisions, while the detailed observations of behavior and descriptions of the course can be better used for improving teaching practices. Further, it is often forgotten that within a particular institution there are usually at least two classes of faculty -- the tenured and the untenured. Evaluation philosophy is likely to differ in the eyes of those two classes if the purposes and assumptions about the consequences of evaluation differ. Even if all members of the faculty participate in the same evaluation system, the ultimate consequences of a poor or an outstanding evaluation are likely to be substantially different for tenured and untenured members. Hence, there are differences in perceived purposes. Sorting out these issues and the kind of data

needed illustrate why an early concern in revising the evaluation process must have to do with defining the purposes of the evaluation.

The question of whether a single faculty evaluation system can serve both personnel decision-making and faculty development needs to be resolved. One current view is that it is necessary to separate the two functions into separate evaluation processes. While most would agree that in some ideal conditions both functions ought to be combined and be mutually supportive, such combination can be very difficult to accomplish because of the different purposes being served. For personnel decisions comparisons are often made, usually across campus and usually using comparative data to help make difficult decisions that result in significant rewards or punishments. A primary goal of such a system is to be perceived as fair and legally defensible. Evaluation for development is more frequently based on measuring change in individual performance and is most often voluntary rather than compulsory. Change is often most rapid when a faculty member discovers that something he or she thought was going well turns out to be perceived by others as an area needing improvement. To discover this requires taking risks in evaluation. Assessment for development needs to encourage experimentation and be tolerant of failure. If the same evaluation system and the same instruments are used for both personnel decision-making and development, voluntary risk-taking and experimentation are less likely to be forthcoming.

Admitting these practical difficulties, it is sometimes argued that the separation of personnel decision-making from development weakens both activities. Even if it is not possible to bring both functions together in a single, integrated system, it may be desirable to bring

them closer together. Will faculty pursue development activities in a system totally segregated from the personnel reward structure? Though faculty probably improve their performance primarily because of an intrinsic drive for excellence and professional pride, external reinforcement for improvement must not be overlooked. Some claim that in a program of faculty development separated from the reward structure the only teachers who participate are those who need development least. There are ways to utilize developmental evaluation during lengthy periods (two-to-three years?) of professional growth and then review those evaluations for personnel decision-making only when key decisions must be made, thus meshing the two systems fairly effectively. Different people in the faculty and administration may be concentrating their efforts on development or the personnel system, but they should have means of communicating with and supporting each other while still maintaining a sense of fairness in the process.

Beyond the basic functions of personnel decisions and professional improvement, faculty evaluation may serve additional explicit or implicit goals. Accrediting agencies often require explicit forms of evaluation. Students want to have influence in evaluating teachers, and they want information about courses and teaching styles that go beyond catalog descriptions. Sometimes an evaluation is used just to provide a facade of objective accountability.

The next exercise should help the study group clarify the single or multiple purposes evaluation is to fulfill. The philosophy of evaluation discussed in the previous chapter should provide the basic approach to serving these varied purposes, but it does not necessarily determine which functions will be served.

FORM 3

Purposes of Faculty Evaluation

Choose one or more of the following options. If more than one option is chosen, rank your choices in terms of primacy of purpose.

1. To provide comparative data about faculty performance upon which personnel decisions will be made. _____

2. To provide data on an individual faculty member's performance (not compared to others) upon which personnel decisions will be made. _____

3. To provide comparative data about faculty performance which will provide a basis to motivate individuals to improve their performance. _____

4. To provide data on an individual faculty member's performance (not compared to others) from which individual faculty members will be motivated to improve. _____

5. To establish base-line data about an individual's performance, against which the individual's growth and development will be evaluated for personnel decisions. _____

6. To provide information for college-wide faculty development efforts. _____

7. To give students a means of registering their views on faculty performance. _____

8. To provide students with consumer information for course choice. _____

9. To conform with accreditation requirements. _____

10. To inform the institutional governing board of the quality of faculty performance. _____

11. To provide data to move from an arbitrary or intuitive personnel system to an information-based system. _____

12. To eliminate undesirable or unproductive faculty members. _____

13. To reward and advance the most desirable and productive faculty members. _____

14. Other purpose: (fill in your own)

V.

IDENTIFYING AND WEIGHTING FACULTY ACTIVITIES

Is every faculty member expected to perform every facet of faculty work equally well? Is every faculty member doing the same things, spending equal time and effort doing them? Not likely, nor is it likely that any institution would wish it so.

Faculty activities most often formally evaluated are teaching and, for most colleges, scholarly or creative production. The faculty perform a variety of additional functions which are expected by the college, but seldom are part of a formal evaluation system. Thus, it should be no wonder that faculty spend little time advising or contribute little to a committee in preference for other kinds of activities. If the college places a value on faculty activities in addition to teaching, those activities should be evaluated and rewarded.

Most evaluation systems assign relative weights to different activities, though often only implicitly and sometimes in esoteric and dysfunctional ways. Evaluation may be improved by 1) making the relative value or weight of key activities (such as teaching, research and campus service) explicit, 2) allowing differential weighting of activities to maximize rewards for peculiar strengths among faculty, and 3) allowing a variety of activities to be evaluated and rewarded beyond the basic core. Such variations and refinements are possible no matter which philosophy of evaluation is followed, though the "summative comparison" approach, if extreme, will limit the freedom to use differential weighting.

Form 4 provides an opportunity to identify and weight faculty activities to be evaluated.

FORM 4

Faculty Activities to be Evaluated and Weighted

The determination about which activities are rewarded can be made college-wide, or by department or by individual negotiation. The important factor is that the faculty understand and accept the list and the relative value placed on different activities. Below is a list of activities in which faculty may engage and for which they may be rewarded.

First, check those activities which are/should be evaluated in your college, then place a percentage rating (or range or percentages) on those you checked. Once agreement is reached on this phase, you have the beginning of an explicit, weighted evaluation system.

	Is Rewarded	Should be Rewarded	Weighting
1. Teaching -- general	_____	_____	_____
a. Pedagogic methods	_____	_____	_____
b. Teaching skill	_____	_____	_____
c. Subject-matter competence	_____	_____	_____
d. Subject-matter difficulty	_____	_____	_____
e. Grading practices	_____	_____	_____
f. Student learning outcomes	_____	_____	_____
g. Student satisfaction	_____	_____	_____
h. Enrollment in courses	_____	_____	_____
i. Other aspects _____	_____	_____	_____
j. Other aspects _____	_____	_____	_____
2. Advising -- general	_____	_____	_____
a. Number of advisees	_____	_____	_____
b. Advisee retention	_____	_____	_____
c. Advisee job and graduate placement	_____	_____	_____
d. Advisee satisfaction	_____	_____	_____
e. Other _____	_____	_____	_____
f. Other _____	_____	_____	_____

		Is Rewarded	Should be Rewarded	Weighting
3.	College service -- general	_____	_____	_____
	a. Assisting other faculty	_____	_____	_____
	b. Effective participation in committees	_____	_____	_____
	c. Assisting in student recruitment	_____	_____	_____
	d. Assisting in fund raising	_____	_____	_____
	e. Success in grantsmanship	_____	_____	_____
	f. Relationships with colleagues	_____	_____	_____
	g. Other _____	_____	_____	_____
	h. Other _____	_____	_____	_____
4.	Community service -- general	_____	_____	_____
	a. Effective participation in community affairs	_____	_____	_____
	b. Community leadership	_____	_____	_____
	c. Other _____	_____	_____	_____
5.	Professional growth -- general	_____	_____	_____
	a. Degree completion	_____	_____	_____
	b. Additional study	_____	_____	_____
	c. Professional publications/ research/creative work	_____	_____	_____
	d. Participation in professional meetings	_____	_____	_____
	e. Leadership in professional associations	_____	_____	_____
	f. Participation in curriculum/ teaching institutes, workshops	_____	_____	_____
	g. Design of new courses/ curriculum	_____	_____	_____

	Is Rewarded	Should be Rewarded	Weighting
h. Use of new methodologies in teaching, advising or research	_____	_____	_____
i. Other _____	_____	_____	_____
j. Other _____	_____	_____	_____
6. Administrative duties -- general	_____	_____	_____
a. Committee leadership	_____	_____	_____
b. Part-time administrative assignments	_____	_____	_____
c. Department/divisional chair	_____	_____	_____
d. Other _____	_____	_____	_____
e. Other _____	_____	_____	_____
7. Personal traits -- general	_____	_____	_____
a. Creativity	_____	_____	_____
b. Friendliness	_____	_____	_____
c. Religious/moral commitment	_____	_____	_____
d. Cooperativeness	_____	_____	_____
e. Intelligence	_____	_____	_____
f. Loyalty to institution	_____	_____	_____
g. Other _____	_____	_____	_____
h. Other _____	_____	_____	_____
8. Miscellaneous	_____	_____	_____
a. Competing job offers	_____	_____	_____
b. Length of service and rank	_____	_____	_____
c. Other _____	_____	_____	_____
d. Other _____	_____	_____	_____

VI.

USING DATA TO SUPPORT EVALUATION

Once activities to be evaluated have been identified and weighted, it is possible to suggest types of data which might be collected. Generally, evaluation will be more valid and fair if more than one kind of evidence is used. The only caveat to be noted about using many sources of information is that the data must be limited to that which can be easily collected and analyzed. Make the system practical or it will soon break down.

Preliminary to the various forms and systems of gathering observations and judgments, there is a need for basic documentation for evaluation. You might consider a kind of archive of regularly collected faculty-supplied or administration-supplied evidence, such as self-reports of activity, syllabi, tests and course materials, copies of publications, professional programs and educational and honorific documents, course enrollment data, advisee assignment data, placement records of students and so on. Much of this "raw" data might be routinely filed at convenient intervals well ahead of the scheduled times for gathering others' views and analyzing the corpus of evidence. The file of such data should be open to review by the faculty member.

Since one common denominator among faculty is teaching, it is probable that one type of evidence will be some kind of student evaluation of teaching. Instruments for this kind of appraisal abound, many of them useful in a variety of colleges. With the voluminous research on the validity and reliability of student evaluations, there can be little excuse for not using well-designed student ratings (for more on student

appraisals, see Appendix 1). But, using only student evaluations for evaluating teaching is equally inexcusable. Students are the most reliable source of information about a teacher's behavior and their reactions to it; they are in a poor position to evaluate course content, the professor's growth over time or other professional activities.

The opinions and observations of colleagues, in the long run, are often the most critical source of data for personnel decisions. Observation especially can also be useful for improving performance. But unlike student rating forms, peer ratings have not enjoyed as much attention from psychometricians or other experts in the evaluation field. And the formal means of recording data collected from colleagues and administrators are likely to more idiosyncratically relate to a particular college and stage of evaluative system development than our student ratings. It is useful to examine peer rating instruments from other institutions, but you will probably need to construct your own.

Form 5, which follows, identifies eight types of data which might be used to support faculty evaluation. Questionnaires/rating instruments are usually some form of student or alumni or colleague evaluation of teaching or advising. Self reports are the faculty member's own reflection on her progress and activities. This type of data becomes especially important for the faculty member to have the opportunity to explain her educational goals or any special circumstances about her performance. Direct observation refers primarily to classroom performance observation, but could also include committee work, advising or other faculty activities. The observation could be by a colleague in a team-taught course, a student, the dean or other colleagues. Interviews of students, colleagues, others can provide more detailed and specific data on an

individual's performance and impact. Expert opinion from outsiders is used primarily for judging a faculty member's competence within his discipline, although it could also be used in other areas. Activity reports are ways for faculty to document what they are doing or how they spend time, such as time studies. This data becomes crucial for faculty whose activities are different from the norm, such as those with heavy advising, committee or project loads. Products can include syllabi, tests, publications, papers, committee minutes, art work -- any product resulting from the individual's professional life. Institutional records are current and historical data about those aspects of the institution which directly relate to the faculty member's work, such as enrollment figures for the individual's classes, attrition rates for his advisees, student exam performance, or placement history for his students.

Note that several copies of Form 5 are provided to cover each type of faculty activity to be evaluated, such as advising, teaching, service or scholarly production. You may need to reproduce copies if your list of activities is extensive.

FORM 5

Data to Support Faculty Evaluation

Use one of these forms for each principal type of activity which will be evaluated (such as teaching, research, service, etc.). Specify which types of data are appropriate for evaluating each activity, then check off which will be required (presumably a smaller number than those types of data which might be appropriate), and finally, indicate who is to supply the data (students in courses, advisees, the person being evaluated, colleagues, department head, dean, alumni, outside experts, etc.).

Activity to be evaluated: _____

Types of Data	Appropriate	Required	Source of Data
Questionnaires/rating instruments (specify by students, colleagues, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
Self-reports	_____	_____	_____
Direct observation of the activity			
_____	_____	_____	_____
_____	_____	_____	_____
Interviews of students, colleagues, others			
_____	_____	_____	_____
_____	_____	_____	_____
Expert opinion from outsiders	_____	_____	_____
Activity reports (specify)			
_____	_____	_____	_____
_____	_____	_____	_____
Institutional records (enrollment figures, attrition reports, placement records, student exam performance, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

FORM 5

Data to Support Faculty Evaluation

Use one of these forms for each principal type of activity which will be evaluated (such as teaching, research, service, etc.). Specify which types of data are appropriate for evaluating each activity, then check off which will be required (presumably a smaller number than those types of data which might be appropriate), and finally, indicate who is to supply the data (students in courses, advisees, the person being evaluated, colleagues, department head, dean, alumni, outside experts, etc.).

Activity to be evaluated: _____

Types of Data	Appropriate	Required	Source of Data
Questionnaires/rating instruments (specify by students, colleagues, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
Self-reports	_____	_____	_____
Direct observation of the activity			
_____	_____	_____	_____
_____	_____	_____	_____
Interviews of students, colleagues, others			
_____	_____	_____	_____
_____	_____	_____	_____
Expert opinion from outsiders	_____	_____	_____
Activity reports (specify)			
_____	_____	_____	_____
_____	_____	_____	_____
Institutional records (enrollment figures, attrition reports, placement records, student exam performance, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

FORM 5

Data to Support Faculty Evaluation

Use one of these forms for each principal type of activity which will be evaluated (such as teaching, research, service, etc.). Specify which types of data are appropriate for evaluating each activity, then check off which will be required (presumably a smaller number than those types of data which might be appropriate), and finally, indicate who is to supply the data (students in courses, advisees, the person being evaluated, colleagues, department head, dean, alumni, outside experts, etc.).

Activity to be evaluated: _____

Types of Data	Appropriate	Required	Source of Data
Questionnaires/rating instruments (specify by students, colleagues, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
Self-reports	_____	_____	_____
Direct observation of the activity			
_____	_____	_____	_____
_____	_____	_____	_____
Interviews of students, colleagues, others			
_____	_____	_____	_____
_____	_____	_____	_____
Expert opinion from outsiders	_____	_____	_____
Activity reports (specify)			
_____	_____	_____	_____
_____	_____	_____	_____
Institutional records (enrollment figures, attrition reports, placement records, student exam performance, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

FORM 5

Data to Support Faculty Evaluation

Use one of these forms for each principal type of activity which will be evaluated (such as teaching, research, service, etc.). Specify which types of data are appropriate for evaluating each activity, then check off which will be required (presumably a smaller number than those types of data which might be appropriate), and finally, indicate who is to supply the data (students in courses, advisees, the person being evaluated, colleagues, department head, dean, alumni, outside experts, etc.).

Activity to be evaluated: _____

Types of Data	Appropriate	Required	Source of Data
Questionnaires/rating instruments (specify by students, colleagues, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____
Self-reports	_____	_____	_____
Direct observation of the activity			
_____	_____	_____	_____
_____	_____	_____	_____
Interviews of students, colleagues, others			
_____	_____	_____	_____
_____	_____	_____	_____
Expert opinion from outsiders	_____	_____	_____
Activity reports (specify)			
_____	_____	_____	_____
_____	_____	_____	_____
Institutional records (enrollment figures, attrition reports, placement records, student exam performance, etc.)			
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

VII.

SCHEDULE AND RESPONSIBILITY FOR COLLECTING EVALUATION DATA

Who is going to collect what data, when and how? Decisions about these issues can make or break the best evaluation design. If confidentiality and fairness are important (and they always are), the responsibility for practical, well-understood and reasonable means of collecting evidence must be carefully planned, assigned and executed.

Forms 6A and 6B are to be utilized for detailing the responsibility and calendar for data collection. Again, there are copies of Form 6A for each major faculty activity. The important point here is that thoroughness must be matched by practical considerations of time and energy, and the natural rhythms of the academic year. Form 6B illustrates that it is important to review these materials to make sure that the burden of work is well-distributed and well-timed.

Later, when producing instructions for each part of the evaluation system, be sure to spell out what is being measured or collected by whom for what purpose, to whom it is being submitted and how the data will be stored, used and how long it will be kept. These specifications are not only necessary under the growing pressure of concern about privacy and protection of individual rights in the storage of personal information; they are simply fair to the people being evaluated and to those supplying data.

FORM 6A

Responsibility for Data Collection

Precisely how does the evaluation system work? Two major details of implementing an evaluation system include assigning responsibility for collecting the data and designating collection times. For each faculty activity to be evaluated, fill out this form listing the required data (assume also that faculty may want to add supplemental data), indicate who is responsible for collecting and reporting this information, when it will be collected and to whom the data is sent.

EXAMPLE

Activity: Advising

Data	Collected By	When	To Whom Sent
Advisee evaluation forms	Advisor	May	Department head
Retention rates of advisees	Registrar	January & September	Department head and dean
Reports from department heads	Department head	May	Dean

Activity: _____

Data	Collected By	When	To Whom Sent

Complete this form for as many major activities as you plan to evaluate.

FORM 6A

Responsibility for Data Collection

Precisely how does the evaluation system work? Two major details of implementing an evaluation system include assigning responsibility for collecting the data and designating collection times. For each faculty activity to be evaluated, fill out this form listing the required data (assume also that faculty may want to add supplemental data), indicate who is responsible for collecting and reporting this information, when it will be collected and to whom the data is sent.

EXAMPLE

Activity: Advising

Data	Collected By	When	To Whom Sent
Advisee evaluation forms	Advisor	May	Department head
Retention rates of advisees	Registrar	Januar & September	Department head and dean
Reports from department heads	Department head	May	Dean

Activity: _____

Data	Collected By	When	To Whom Sent

Complete this form for as many major activities as you plan to evaluate.

FORM 6A

Responsibility for Data Collection

Precisely how does the evaluation system work? Two major details of implementing an evaluation system include assigning responsibility for collecting the data and designating collection times. For each faculty activity to be evaluated, fill out this form listing the required data (assume also that faculty may want to add supplemental data), indicate who is responsible for collecting and reporting this information, when it will be collected and to whom the data is sent.

EXAMPLE

Activity: Advising

Data	Collected By	When	To Whom Sent
Advisee evaluation forms	Advisor	May	Department head
Retention rates of advisees	Registrar	January & September	Department head and dean
Reports from department heads	Department head	May	Dean

Activity: _____

Data	Collected By	When	To Whom Sent

Complete this form for many major activities as you plan to evaluate.

FORM 6A

Responsibility for Data Collection

Precisely how does the evaluation system work? Two major details of implementing an evaluation system include assigning responsibility for collecting the data and designating collection times. For each faculty activity to be evaluated, fill out this form listing the required data (assume also that faculty may want to add supplemental data), indicate who is responsible for collecting and reporting this information, when it will be collected and to whom the data is sent.

EXAMPLE

Activity: Advising

Data	Collected By	When	To Whom Sent
Advisee evaluation forms	Advisor	May	Department head
Retention rates of advisees	Registrar	January & September	Department head and dean
Reports from department heads	Department head	May	Dean

Activity: _____

Data	Collected By	When	To Whom Sent

Complete this form for as many major activities as you plan to evaluate.

FORM 6B

Annual Calendar of Evaluation Activity

Making an annual calendar of activities will provide a display of the entire evaluation system and an easy review of the work flow in the process. Devise the data for this form from your complete set of forms 6A. After you have displayed all the information, you may want to revise this form and some of 6A to achieve a better distribution of work among people and the times of the year.

One additional activity is also included in this form: Analysis of the data and action. Do not underestimate the amount of time taken at this stage. For instance, if an all-college committee is charged with evaluating all faculty members up for tenure and promotion, even in a small college they may have to review 25 or 30 people. If a significant amount of data has been collected, this process consumes an enormous amount of time. Be generous in the amount of time allowed.

When	Data	Collected By	To Whom Sent	Time of Analysis & Decisions

VIII.

INSTRUMENTS AND PROCEDURES, ADOPTION, IMPLEMENTATION, EVALUATION AND FURTHER REVIEW

If all of the previous forms have been completed to your satisfaction, you now have the basic structure of an evaluation system. Your probable next steps are to design specific instruments and office procedures to implement the system. These details should be worked out carefully before the system is proposed for adoption. Such a task is best done by only a few people who will present the forms and procedures to the study group for review and revision. If, for instance, the procedure requires the review of a variety of student rating instruments, it may take some time to collect the various instruments and assess their applicability to your situation. However, some of this activity can take place after the basic system has been adopted if time is a critical factor. The important thing to remember is not to begin using any form or procedure before it has been formally reviewed and adopted by the critical decision-making body -- usually the full faculty. We have provided no work form to list instruments and procedures, because the variations are infinite. But the study group should be satisfied that all details are listed and completed.

Adoption. Be sure to carefully review the entire study process outlined on Form 1 just prior to adoption to be sure that all the appropriate steps have been followed. If the process has been open, suggestions incorporated into the system and sufficient time allowed for discussion, there should be little problem having the new system adopted.

Implementation. The study group might helpfully review implementation details with those who are to administer the system. The first time

through may require considerable modification and adjustment. It is important to remember to inform the faculty generally as changes in the system are being made, even if they are relatively minor. "No surprises" is a good motto for the first year of a new evaluation system.

Evaluation and review. As soon as the evaluation system has been run through one complete cycle, some group should report back to the entire faculty and appropriate others on how well it has worked. You may want to give some consideration to how this evaluation of the evaluation system is to be designed before you begin using the system for the first time. The design of this study could be very complex if resources are available and you need to know how newly-designed forms and processes are working, whether they are reliable, and so forth. But you may only want to know whether or not faculty, students and administrators are satisfied with the system. In either case, you need to be explicit about what the criteria for success are.

If the new system has been well designed and is clearly acceptable, it may run without substantial revision for several years. It is more likely to run smoothly, however, if a major study is undertaken after two or three years of use. Even if the conclusion of this review is that only minor modifications are necessary, it is well worth doing. No system will last forever. Part of the original legislation adopting a system might very well include a mandated review at some future date. This will reassure everyone involved that everything, even the evaluation system, is to be judged on the basis of merit and not simply administrative expediency.

GENERAL BIBLIOGRAPHY

Aleamoni, Lawrence M. "Proposed System for Rewarding and Improving Instructional Effectiveness." College and University, 51 (Spring 1976), 330-338.

Centra, John. Renewing and Evaluating Teaching. San Francisco: Jossey-Bass, 1977.

Doi, James I., Ed. Assessing Faculty Effort, Vol. II of New Directions for Institutional Research. San Francisco: Jossey-Bass, 1974.

Eble, Kenneth E. The Recognition and Evaluation of Teaching. Washington, D.C.: American Association of University Professors, 1971.

Faculty Evaluation for Improved Learning. Atlanta: Southern Regional Education Board, 1977.

Grasha, Anthony. Assessing and Developing Faculty Performance. Cincinnati, Ohio: Communication and Education Associates, 1977.

Kulik, James A. and McKeachie, Wilburt J. "The Evaluation of Teachers in Higher Education," in F.N. Kerlinger, Ed., Review of Research in Education, Vol. III. Itaska, Illinois: F.E. Peacock Publishers, Inc., 1975.

McIntyre, Charles. "Evaluation of College Teachers." Criteria, No. VI. Ann Arbor, Michigan: Center for Research on Learning and Teaching, University of Michigan, 1977.

Miller, Richard I. Developing Programs for Faculty Evaluation. San Francisco: Jossey-Bass, 1973.

Miller, Richard I. Evaluating Faculty Performance. San Francisco: Jossey-Bass, 1972.

Romney, Leonard C., and Manning, Charles W. Faculty Activity Analysis: Interpretation and Uses of Data. Denver: National Center for Higher Education Management Systems Technical Report 54, 1974.

Scholl, Stephen C. "Evaluation of Teaching," in Scholl and Inglis, Sandra C., Eds. Teaching in Higher Education: Readings for Faculty. Columbus, Ohio: Ohio Board of Regents, 1977.

Seldin, Peter. How Colleges Evaluate Professors. Croton-on-Hudson, New York: Blythe-Pennington, Ltd., 1975.

Smock, H. Richard and Crooks, Terrance J. "A Plan for the Comprehensive Evaluation of College Teaching," Journal of Higher Education, 44 (November, 1973), 577-586.

Yuker, Harold E. Faculty Workload: Facts, Myths and Commentary. Washington, D.C.: ERIC/Higher Education Research Report No. VI, 1974.

APPENDIX 1

USING STUDENT APPRAISAL INSTRUMENTS

Too often the major focus of discussions about faculty evaluation is the use of student evaluations or appraisal of instruction. It is hard to imagine a college or university that does not utilize some kind of student appraisal of instruction as an essential part of a faculty evaluation system. But these instruments are only one source of information, and it is unfortunate that their discussion is so highly charged and political. One good result of all this conversation and concern, however, is that there has probably been more research and development of student rating instruments in recent years than has been lavished on any other facet of faculty evaluation.

The questions most often asked and researched regarding student rating instruments are: "Are these instruments reliable?" and "Are they valid?" In brief, various investigators have found that many instruments are reliable (that is, they are internally consistent and the results achieved are stable over time -- or, they accurately measure the same thing each time they are used). The validity issue, on the other hand, may never be settled. Validity is a measure of whether or not an instrument is evaluating what it is supposed to evaluate. If there is little agreement between researchers and users about what is supposed to be measured by student rating instruments, it is difficult to produce a "valid" instrument.

Most people who favor the use of student ratings as an important component of an evaluation system claim that they are intended to measure

student reactions to instruction, and are therefore intrinsically valid.[1] Most frequently, those who oppose the regular use of such instruments argue that they cannot be valid because the evidence for effective teaching is what students learn, not what students think about what they are learning or about what the teacher is doing.[2] We come down on the side of those who claim that it is important to know how students feel about what they are learning, and to have their reports on what teachers are doing when they teach. It is also possible to know what students themselves are doing by the use of student ratings. We openly admit that student rating systems are not valid measures of how much students learn. In fact, there is scant evidence that we can satisfactorily and easily measure what students learn in a way that is acceptable to everyone. That dilemma is one of the major challenges facing educational researchers today.

Assuming that some kind of student rating instrument will be used, the following general findings about rating instruments might be helpful.

1. Students are honest and say what they believe (as long as the instrument is administered in a fair and confidential manner).
2. Students are telling you what they perceive. Their perceptions are as real as yours are, even though they may differ.

---

[1] Wilbert J. McKeachie, "Student Ratings of Instruction," AAUP Bulletin, 55 (1969), 439-444.

[2] F.N. Kerlinger, "Student Evaluation of University Professors," School and Society, 99 (1971), 353-356.

3. Rating instruments tell you whether the instruction being delivered meets students' perceived needs and meshes with their individual learning styles. The more specific the question, the better the information. For example, if a student is asked to "tell me three things to do to improve my instruction," he or she is likely to tell you some very specific things that will be helpful to meet that student's needs. Those things might not be applicable to other students.

4. There is some positive weighting, or "halo" effect in ratings associated with the age and rank of professors (older, tenured professors tend to receive higher ratings). Those who are experimenting with new approaches to teaching (such as a Rogerian discussion technique, or a radically student-centered course design) can expect more negative ratings. Students become anxious if they are on unfamiliar ground. Such ratings should not, of course, discourage experimentation.

5. There are some items on which the professors generally rate so highly that they provide little discrimination among most instructors. These include statements indicating that the teacher is friendly, enjoys teaching, has a thorough knowledge of the course, seems knowledgeable, is enthusiastic, maintains good feelings in the class, likes students to disagree and ask questions and gives individual attention to students. When someone does not score high on such items, we should pay attention to these areas. Other questions are likely to receive relatively low ratings from most students for most instructors, including questions relating to the use of audiovisual materials, motivation to take additional related courses,

the value of laboratory sessions, the interest and stimulation of reading assignments and textbooks, and the student's sense that he or she is performing up to his or her potential or has gained a better understanding of himself/herself through a particular course. Hence low ratings here are seldom a cause for alarm.[3]

6. Students tend to give helpful information that does differentiate between professors on questions that regard general teaching skill, classroom interaction, the amount of feedback given to students, the way discussions are conducted, classroom organization, preparation, and whether or not the teacher requires a significant amount of work for the course. Interestingly, it is usually found that students do not reward a professor because he or she is "easy," except when the student already has received a grade that is very different from the one the student expected.
7. Students do not generally rate the professor in direct correlation to the grade the student expects to receive in the course. Often, in fact, students who do very well are harder on their professors than students who do poorly.
8. The effects of class size, required and not required courses, disciplinary biases and the like can be accommodated by a sensitive analysis of rating results. The most common advice is to compare instructors who teach in similar circumstances with each other. For

---

[3] James Kulik, "Early Responses from The Campus Use of The Instructor -- Designed Questionnaire (IDQ)," in S.C. Ericksen, "The Lecture," Memo to the Faculty, 60 (April, 1978), 5.

instance, it is unfair to compare the ratings of a senior professor of microbiology teaching a research seminar with a junior professor teaching a required freshman English course.

9. There is some evidence that student evaluation prompts improvement in teaching, especially when a professor receives ratings that are lower than he expected on specific items.

The process of choosing a student appraisal instrument can provide a fine occasion for discussions about teaching. Since there are many well-designed, proven instruments available, it is wise to consider the benefits of adopting one of these rather than creating a new instrument for your own campus. Designing good instruments is very time consuming and may not be worth the effort as compared with a reasonably good system that is ready to run.

Among computer-based systems, there seems to be greater acceptance by faculty of those instruments which allow individual professors to play some role in customizing the questionnaire, while maintaining a core of common questions that can be compared across departments and disciplines (see, for example, CAFETERIA, IDEA and IAS below).

Below are brief descriptions of some widely used instruments. They illustrate some options currently available.

1. "CAFETERIA" type instruments.

Several years ago the Measurement and Research Center at Purdue University developed a course appraisal system called CAFETERIA. This instrument allows a faculty member to construct an individualized instructor and course appraisal consisting of five "core" items used by all instructors and an additional 35 items

chosen from a catalog of 200 questions concerning a variety of course and instructor issues. The results are analyzed by computer, including responses to a few "instructor-supplied items" on which no norms are available. A norm base has been developed for Purdue University and also for a set of Indiana Independent Colleges that provides for comparative percentile scores that are printed on each professor's report in addition to the median responses to each item. For more information write Dr. James Derry.

Today there are several other CAFETERIA-type systems in use. The complete CAFETERIA package can be purchased from the Purdue Research Foundation (ready to be mounted on any computer that can accept a contemporary version of FORTRAN). Purdue will also process instruments for some colleges. A "Cadillac" version of CAFETERIA is available from the Office of Instructional Resources at the University of Illinois at Champaign-Urbana 61801: the Instructor and Course Evaluation System (ICES) allows a variety of reporting formats and information for the instructor, a catalog of over 600 items to choose from, and more sophisticated reporting mechanisms. This system uses only three "core" items: "rate the course content," "rate the instructor," and "rate the course in general". Special items are indicated for teaching situations such as science laboratories or field trips, and for various interests the teacher may have, ranging from warmth and concern for students to the effective use of computers.

A more simplified version of the CAFETERIA type is the Instructor-Designed Questionnaire (IDQ), available for modification by any college through the Center for Research on Learning and Teaching

(CRLT), University of Michigan. A description of this system was published as a Memo to the Faculty by CRLT in 1976, entitled "Student Reactions to Instruction." That six-page monograph includes the complete catalog of 153 items that is similar to the one from Purdue but available without charge and arranged in a somewhat more effective manner.

The CAFETERIA approach is used at most of the "Big 10" universities and a new version for small colleges is being developed by the Great Lakes Colleges Association Faculty Development Program. All of the CAFETERIA variations are most useful for developmental purposes, rather than for summative comparative judgments for personnel decisions. However, if they are used consistently, over a long period of time, the "core" items can form a useful basis for personnel judgments if sensitively analyzed.

2. Instructional Development and Effectiveness Assessment (IDEA).

In recent years the IDEA system has been tried by many colleges across the country as a result of a large Kellogg Foundation grant for dissemination administered by the Center for Faculty Evaluation and Development at Kansas State University. This set of instruments requires that the professor identify course and learning objectives and then the students rate their progress toward these objectives. The professors are given suggestions about areas of emphasis which would raise students' responses to various items in the analysis report. Cost of implementing IDEA is relatively high, but services from Kansas State University are extensive. Information may be obtained from the Center for Faculty Evaluation and Development, Post Office Box 3000, Manhattan, Kansas 66502.

3. Student Instructional Report (SIR).

This relatively short instrument is the oldest available commercial student appraisal. It asks questions about the course pertaining to teacher-student relationships, course objectives and organization, lectures, reading assignments, course difficulty and examinations. Additionally, students are asked about their own motivation for taking the course and expected grade. Like the IDEA system, cost is relatively high, but the Educational Testing Service provides excellent processing services for a campus not willing or equipped to set up its own system. Available from the Educational Testing Service, Princeton, New Jersey 08540.

4. Teaching Analysis By Students (TABS).

TABS is part of a more comprehensive teaching improvement consultation process that was developed at the former Clinic to Improve University Teaching at the University of Massachusetts. The instrument is particularly useful for classroom diagnosis and follow-up measurement of improvement. A college intending to use TABS should seek consultation from someone familiar with the "clinic" process to gain its full benefits. This system is described fully in the Bergquist and Phillips Handbook for Faculty Development, Volume II.

5. Instructional Assessment System (IAS).

This set of instruments is described in the Bergquist and Phillips Handbook for Faculty Development, Volume I. It is a collection of seven instruments for different teaching situations and raters, for example: self-evaluation, peer evaluation, lecture courses, studio courses. The materials may be copied directly from the handbook. No scoring service is available.

6. Course Evaluation Questionnaire (CEQ).

One of the oldest well-researched computerized instruments, CEQ contains 23 "general concept" items about course attitude, instructional method, course content, student interest and the instructor. It may be used, now, as part of ICES, described above. Services for using this system are available and more information may be obtained from the Office of Instructional Resources, Measurement and Research Division, 307 Engineering Hall, University of Illinois, Urbana, Illinois 61801.

While most of the instruments described above are relatively easy to install on a computer or score by hand, and they are all rather good, they are not without problems. Often the campus reaction is negative because the instrument is simply installed without going through the rigorous process of adoption and particularization that other elements of an evaluation process have suffered. However, many colleges have inadequate time or resources to develop their own forms, have too few instructors and opportunity to establish reliable norms, and thus should consider the use of a pre-packaged system. Simply remember: campus review of the various systems and debate is essential.

One other caveat: complex rating instruments, while popular, are usually not the most helpful means of gathering data to change a course plan or make a specific improvement in teaching. For instance, asking students directly, or in writing, at midterm: "What can I do to improve this course next week" may be much more useful (if you follow through).

For a fuller discussion of the issues involving student appraisal instruments, refer to the following sources:

Aleamoni, Lawrence M. and Yimer, M. "An Investigation of the Relationship Between Colleague Rating, Student Rating, Research Productivity, and Academic Rank in Rating Instructional Effectiveness." Journal of Educational Psychology, 1973, 64, 274-277.

Costin, R., Greenough, W.T., and Menges, R.J. "Student Ratings of College Teaching: Reliability, Validity, and Usefulness." Review of Educational Research, 1971, 41, 511-535.

Doyle, Kenneth O., Jr. Student Evaluation of Instruction. Lexington, Massachusetts: Lexington Books, 1975.

Drucker, A.J. and Remmers, H.H. "Do Alumni and Students Differ in their Attitudes Toward Instructors?", Journal of Educational Psychology, 1951, 42, 129-143.

Ericksen, Stanford C. "Student Evaluation of Teaching." Criteria, No. II (June, 1976), Center for Research on Learning and Teaching, University of Michigan.

French-Lazovick, Grace. "Evaluation of College Teaching," Association of American Colleges Occasional Paper, 1976.

Gage, N.L. "Student Ratings of College Teaching: Their Justification and Proper Use." In Allen, Melnik, and Peele, Eds., Reform Renewal Reward. Amherst, Massachusetts: Clinic to Improve University Teaching, University of Massachusetts, 1975, 121-135.

Kulik, James. "Student Reactions to Instruction," Memo to the Faculty. Ann Arbor, Michigan: University of Michigan, 1976.

Landers, Audrey. Student Evaluation of Teaching: Issues and Instruments. University of Alabama Teaching-Learning Center, Box 1443, University, Alabama 35486. 1976.

APPENDIX 2

A CASE FOR DIFFERENTIAL EVALUATION

Differential evaluation allows different units/departments or even individuals to determine the relative importance or weight of various activities to be evaluated. Not only is this approach different from the standard practice of evaluation by a college-wide formula, but it requires thorough consideration of the philosophical implications of allowing individuals to receive similar rewards (such as promotion and tenure) for sometimes significantly different kinds of activity. Using differential evaluation, different departments could weight scholarly activity more heavily than others, or some individual faculty members might be evaluated more heavily for their work in student advisement than would their colleagues. Such an approach allows for individual differences and needs, among departments or among faculty within a department.

In a hypothetical case, South College has identified the following activities as those to be evaluated and established ranges of weight for each:

Teaching --	40-80 percent, average 60 percent
Advising --	0-40 percent, average 20 percent
College Service --	10-20 percent, average 10 percent
Scholarly and Creative Growth --	0-20 percent, average 10 percent
	Total: 100 percent

Another college might include advising as part of teaching or service, give much more weight to scholarship, and consider community service as an additional category. There are obviously many variations on this pattern, and sometimes one department might have a category that no other department in the college has.

Using such a college-wide guideline, a dean or personnel committee would negotiate with each department yearly to identify variations on the college norm necessitated by departmental needs. Department A, for example, is undergoing extensive curriculum revision and presents a case for all of its faculty to be evaluated less for teaching and more for college service, since much energy and time would be devoted to the curricular revision.

Activity	College Norm	Dept. A (All Faculty)	(Prof. A)	Dept. B (Prof. B)	(Prof. C)	Dept. B. Average Total
Teaching	60%	50%	80%	40%	60%	60%
Advising	20%	20%	0%	40%	20%	20%
College Service	10%	20%	10%	20%	0%	10%
Scholarly Growth	10%	10%	10%	0%	20%	10%
Totals:	100%	100%	100%	100%	100%	100%

In the case of Department B the chair wished to take into account individual differences in circumstances among the faculty, so she proposed a differential evaluation among individuals. When averaged, these

differences conformed to the college norm (though this need not necessarily be the case). Professor A teaches well, but his advising skills are so bad that he has been relieved of that responsibility (and he has indicated that sometime during the year he will attend some advising skills workshops). Professor B, an excellent advisor, has her teaching load reduced so that she may handle more advising; she also serves on many committees and is willing to take no credit for scholarly growth this year. Professor C has been relieved of college-wide commitments so that he can complete his dissertation this year. In this hypothetical case, the dean assumed that the college would allow individual variations within the department only if the average total for the department came very close to the college norm so that departments would not veer too sharply from each other in their performance.

Differential evaluation requires that college-wide norms be developed and that deans or department heads negotiate with their faculty and fully consider both individual and departmental needs each year. Such a process is particularly helpful if a college or department begins to experiment with growth contracts (see Appendix 3), but is also a way to work out the natural differences between work assignments among departments that occur in almost every college (such as the differences between the English Department, Physical Education and Music).

APPENDIX 3

GROWTH CONTRACTS

One of the newer approaches to faculty evaluation and development is the growth contract. Examples of institutions already using growth contracts are Gordon College, New College at the University of Alabama, Ottawa University, Los Medanos College, Columbus College and Austin College. Here is a process by which individual faculty assess their professional strengths and those areas they wish to improve. They write a growth contract or plan which is negotiated or shared with colleagues, the department head, the dean or a committee. The contract may include a stipulation of what resources, time or other support the institution will provide the individual to meet the goals. Each year progress on the growth contract is evaluated. The major evaluation usually occurs at a midpoint assessment stage and secondarily at the end of the contract.

Colleges differ widely in the way growth contracts are utilized. Some use them informally as a supplement to summative evaluation; for others it is the only basis for evaluation. Some require negotiations with the department head or dean; others encourage sharing with colleagues. Some employ a complex system of committees for each contract; others use no formal supervision. Others have offered personal growth or life-planning workshops which culminate in writing of growth plans which may or may not be "contractual" with the institution.

The important points about growth contracts are that they are developmental, that they allow the individual to determine what will be evaluated, that they force consideration of evaluation at the beginning of a year or period of years, and that they are relatively easy to

evaluate. It is not difficult to assess whether or not the person did what he or she said was going to be done as opposed to assessing whether or not the faculty member is a good teacher or meritorious member of the community. Major problems with growth contracts are that it is difficult to establish reasonable expectations the first few times and it is difficult to achieve comparability between individuals' growth.

For further information, see:

Buhl, Lance C. and Greenfield, Adele. "Contracting for Professional Development in Academe." Educational Record, 56, (Spring 1975), 111-121.

Gross, Richard, "Faculty Growth Contracts." Educational Horizons (Winter 1976), 74-79.

Hodgkinson, Harold L. "Faculty Reward and Assessment Systems." In B.L. Smith, ed., The Tenure Debate. San Francisco: Jossey-Bass, 1973.